Asunto: The informatics of DNA: Letters, words, sentences, texts, and their meanings
De: Sungchul Ji <sji@pharmacy.rutgers.edu>
Fecha: 27/11/2017 0:02
Para: FIS FIS <fis@listas.unizar.es>

Hi FISers,

We may have in DNA a golden opportunity to define what **information** is.

(1) We now know that we are different from mice because our DNA sequences are different from those of mice [1]. That is, we are different from mice because our DNA carries different kinds (both with respect to *quality* and *quantity*) of INFORMATION from the mouse DNA:

"When it comes to protein-encoding genes, mice are 85% similar to humans. For non-coding genes, it's only about 50%. The National Human Genome Research Institute attributes this similarity to a shared ancestor about 80 million years ago." http://www.thisisinsider.com/comparing-genetic-similarity-between-humans-and-other-things-2016-5

(2) We also know that our properties or behaviors are at least in part determined by both DNA sequences (i.e., *genetics*) and the way they are turned on or off by environment-sensitive cells constituting our body (i.e., *epigenetics*): We are the products of both our *genes* and our *environment*. The causal link between DNA and our behaviors can be briefly summarized as follows:

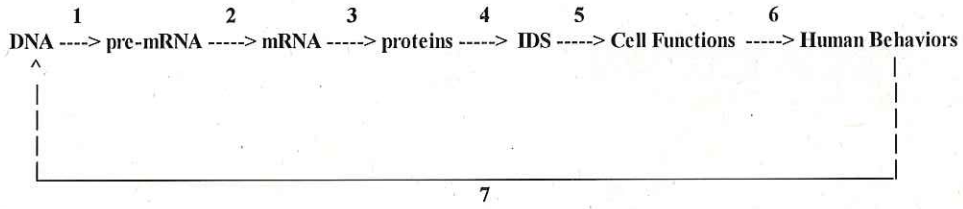


Figure A. The flow of genetic and epigenetic informations between DNA and the human behavior. IDS stands for the *Intracellular Dissipative Structures* (also called the *Dissipative Structures of Prigogine*) such as ion gradients across cell membranes and within the cytoplasm without any membrane barriers. According to the Bhopalator, a molecular model of the living cell proposed in 1985 in a meeting held in Bhopal, India, IDS's are postulated to be the immediate or the proximal causes for all cell functions [2]. The seven steps in the scheme are
1 = transcription
2 = splicing
3 = translation (explained in (3) in more detail.)
4 = enzyme catalysis
5 = cell motions
6 = body motions
7 = the effect of human behavior or emotion on gene expression, e.g., see the phenomenon of the *conservedtranscriptional response to adversity* (CTRA) [3].

I hope that the *information flow scheme* shown in **Figure A** can serve as a concrete example of information inaction as information scientists strive to come up with a generally acceptable definition of what INFORMATIONis.

(3) Unlike in Steps 1 and 2 where the same kinds of molecules, i.e., the nucleic acids, DNA and RNA, directly interact (or contact or touch each other) via the Watson-Crick base-paring mechanism (see the second row in **Figure 1** below), in Step 3, there is no such direct interaction between mRNA and amino acids, but rather their interactions are mediated by tRNA which recognizes mRNA at its *anti-codon arm* and amino acids at its 3′-*acceptor stem*, about 60 angstroms away (see the blue region in the mechanism of translation shown at https://www.quora.com/Why-are-ribosomes-so-important-in-plant-cells). The universality of the wave-particle duality demonstrated in [4] suggest that the tripartite coupling among codon, anticodon, and amino acid in the ribosome-mRNA-tRNA complex may be mediated by *resonant vibrations* or *standing waves* (also called *resonance* or *resonant waves*) generated within the complex, just as the vibratioal patterns located at distant regions on the Chladni (1756-1827) plate [5, 6] are coordinated via resonance.

The Chladni plate [5, 6] is an ideal model for illustrating the role of resonance in molecular biology. At a given resonance frequency, the particles on remote regions of the Chaldni plate are coordinated without any direct interactions between them and yet form ordered patterns. To me this is similar to what happens in the ribosome system when a peptide molecule is synthesized; i.e, different components of the ribosome-mRNA-tRNA complex execute their motions that are so coordinated as to achieve the peptide synthesis. The ribosome and the Chladni plate are compared at several levels in **Table 1.**

Table 1. The ribosome-Chladni plate comparison.

	Chladni plate	Ribosome
1. Vibrating system	metal plate	ribosome + mRNA + tRNA complex
2. Objects being organized	sand particles	codons, anticodons, aminoacyl residues
3. Scale	macroscopic	microscopic (also called molecular)
4. Source of energy driving the organization	electricity/sound waves	thermal energy paid back by the free energy of chemical reactions without violating the Second Law [7]
5. Common principle obeyed	Fourier theorem	Fourier theorem

The idea of using the Chladni plate as a molecular model of the protein synthesis in the ribosome occurred to me while preparing for the keynote lecture, "RASER model of single-molecule enzyme catalysis", that I presented at the **8th International Conference on Proteomics and Bioinformatics**, held in Osaka, Japan, in May, 2017 (see the right bottom panel in **Figure B**). I formulated this idea before I realized that J. Franck and his group was working on the mechanism of protein synthesis in the ribosome along based on a similar approach [8] (i.e., the wave approach to molecular biology [4]), for which he shared the Nobel Prize in Chemistry this year [9].

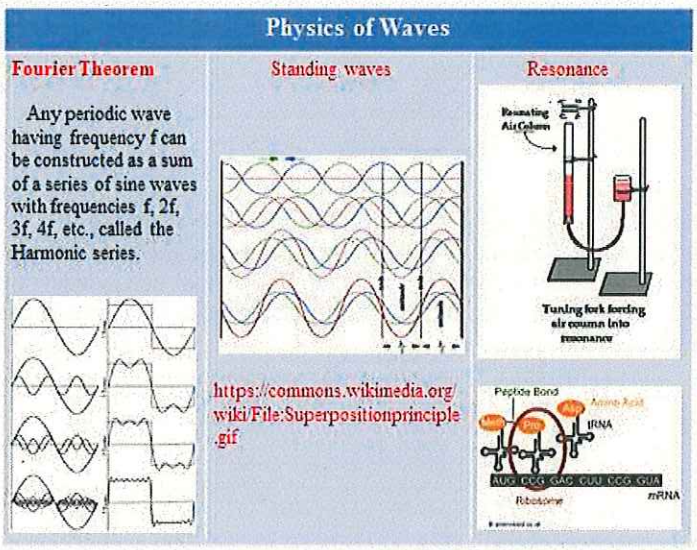


Figure B. *The wave-approach to unraveling the molecular mechanisms underlying protein synthesis in the ribosome*: *The role of resonance*.

There are several components in what is here referred to as the "resonance model of the ribosome structure and function' (RMRSF):
(i) The ribosome and its associated RNA's and proteins are a system of molecular oscillators that obey the Fourier theorem (see the left panel in **Figure B**), i.e., all the wave patterns, both traveling and standing (see the middle panel in **Figure B**) of the system can be constructed from the linear combinations of the vibrational motions of all the chemical bonds [10] ,
(ii) the standing waves in the ribosome system are determined by the geometrical shape (or topology) and the energy content of the system, and
(iii) the chemical (i.e., peptide bond formaton) and mechanical events (i.e., translocation of the tRNA from the A site to the P site, etc.) taking place at the anti-codon arm and the 3'-acceptor stem are 'coupled' or 'coordinated' through the resonance mechanism, just as are the motions of the sand particles located in different regions on the Chladni plate, or just as the vibrational motions of the tuning forks are coupled to the vibrational motions of the water column by resonance thereby exchanging both energy and information between them (see the right panel in **Figure B**).
The resonance model of the *ribosome structure and function* depicted in **Figure B** is based on the simiarilties and differences between the ribosome and the Chaladni plate summarized in **Tale 1**.

It may be worthwhile to point out that my vibrational approach to enzymology/molecular biology dates back to 1974 [11] when I concluded as follows:

" *An ordinary enzyme possesses 10^3 to 10^4 vibrational degress of freedom, as compared to 3 each for the translational and*

rotational degrees of freedom. It is therefore reasonalble to assume that the vibrational motions of individual bonds in the enzyme will be far more important in enzyme catalysis than the translational or rotatioal motions of the enzyme as a whole. Given all the vibrational frequencies of the individual bonds in an enzyme, as well as their three-dimensional arrangements, we can in principle deduce the thermodynamic and catalytic properties of the enzyme under any conditions."

(4) The resonance mechanism of coupling proposed in **Figure B** between the aminoacyl group at the 3' acceptor stem and the anticodon at the opposite side of the ribosome about 60 angstroms away is consistent with the conclusion reached by Petoukhov that *genes in DNA can be viewed as oscillators* that couple to one another via resonance [12]. One of the reasons that Petoukhov came to invoke the *resonance mechanism* is that many of the regularities found in the genetic codes and DNA sequences (e.g., Chargaff's parity rules [13]) obey the *tensor multiplication rules of matrices* widely used in engineering and physics to analyze vibrational mechanics [12]. Thus, it is clear that Petoukhov's resonance approach to modeling genetic structures [12, 13] indirectly supports the resonance coupling mechanism postulated for the transfer of genetic information from the codons of mRNA to the amino acyl residues of tRNA as briefly summarized in **Figure B** above.

(5) The four bases known as the genetic alphabet, i.e., A, C, G, and T (see the first row in **Figure 1**) divide into the well-known Watson-Crick base pairs, i.e., the T-A pair characterized by two hydrogen bonds and the C-G pair characterized by 3 hydrogen bonds (see the second row). In addition to this Watson-Crick pair, there are two other pairs less well known -- the amino (A, C) and keto (G, T) pair, and the pyrimidine (C, T) and purine (A, G) pair, each pair being characterized by unique molecular features or traits (and hence unique vibrational resonance properties as well) (see the third row). These three pairs of the nucleobase pairs (i.e., *pairs of pairs*) can be represented as a triadic set of *four* binary numbers as shown on the right-hand side in the second row of **Figure 2** (omitting the subscripts), namely, 0011, 0101, and 1001. Stated equivalently, each of the four nucleobases can be represented either as 0 or 1 within a given pair of base pairs. That is, G and A can be assigned the same binary digit, 0, and T and C the same binary digit 1, within the pyrimidine-purine pair. On the other hand, the bases within each pair are assigned either 0 or 1. For example, within the 2H-3H pair, G and C are assigned 1 and and A and T are assigned 0. In contrast, within the pyrimidine-purine pair, T and C are assigned 1 and G and A are assigned 0, etc. Two things are noteworthy at this point:

(i) The addition of the first two rows of the binary digits in **Figure 2** produces the third row, if the addition is carried out based on the rule of *modulo-2 addition*: i.e., addition of two identical digits = 1; addition of two different digits = 0. This in effect reduces the three base pairs into two pairs. That is, knowing any two of the 3 yin-yang pairs allows us to infer the third correctly. Nature may have utilized this fact in designing its error-immunity coding in genetics [12, 13].

(ii) Any sequence of nucleotides can carry two overlapping binary messages [12], each message being decodable based on one of the the two alternative base (or yin-yang) pair rules in **Figure 2**. An example is given below:

GCTCGCCTAA

0 1 1 1 0 1 1 1 0 0 according to the *pyrimidine-purine* pair rule

0 1 0 1 0 1 1 0 1 1 according to the *amino-keto* pair rule

In other words, the same set of 10 nucleotides above can encode two independent binary messages, leading to the conclusion that DNA is akin to a *broadband* communication channel, 'broadband' indicating the ability of a communication channel to transfer information fast because of its ability to send multiple messages simultaneously.

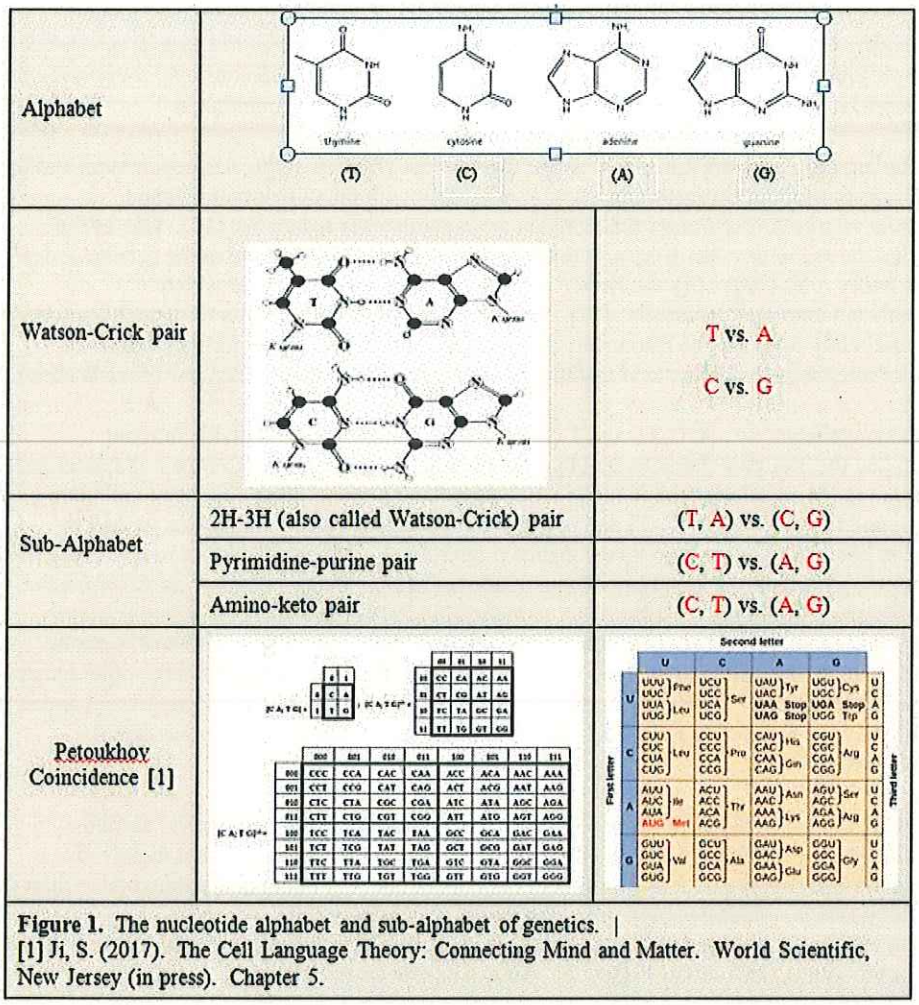
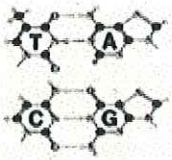


Figure 1. The nucleotide alphabet and sub-alphabet of genetics. |
[1] Ji, S. (2017). The Cell Language Theory: Connecting Mind and Matter. World Scientific, New Jersey (in press). Chapter 5.



TRAITS	G	A	T	C
1) pyrimidine (C,T), purine (A,G)	0_1	0_1	1_1	1_1
2) amino (A,C), keto (G,T)	0_2	1_2	0_2	1_2
3) complementarity (C,G) and (A,T) with 3 or 2 hydrogen bonds	1_3	0_3	0_3	1_3

Fig. 4. Three binary sub-alphabets according to three kinds of binary-opposite traits in a set of nitrogenous bases G, A, T, C. Left: the molecular structure of these bases of DNA. Right: the partition of the four-letter alphabet of DNA on three binary sub-alphabets in accordance with three binary-oppositional traits. Inside each binary sub-alphabet, equivalent letters are marked by the symbol 1 or 0.

Figure 2. The three *yin-yang pairs* [14] (also called *binary opposite pairs* or *binary sub-alphabets* [12, 13]), 1), 2), and 3), that can be generated from the genetic alphabet A, C, G and T.

[C A; T G] =

	0	1
0	C	A
1	T	G

; [C A; T G]^2 =

	00	01	10	11
00	CC	CA	AC	AA
01	CT	CG	AT	AG
10	TC	TA	GC	GA
11	TT	TG	GT	GG

[C A; T G]^3 =

	000	001	010	011	100	101	110	111
000	CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
001	CCT	CCG	CAT	CAG	ACT	ACG	AAT	AAG
010	CTC	CTA	CGC	CGA	ATC	ATA	AGC	AGA
011	CTT	CTG	CGT	CGG	ATT	ATG	AGT	AGG
100	TCC	TCA	TAC	TAA	GCC	GCA	GAC	GAA
101	TCT	TCG	TAT	TAG	GCT	GCG	GAT	GAG
110	TTC	TTA	TGC	TGA	GTC	GTA	GGC	GGA
111	TTT	TTG	TGT	TGG	GTT	GTG	GGT	GGG

Figure 3. The tensor multiplications of 2x2 matrices, e.g., [C A; T G], to generate a 4x4 and an 8x8 matrices. That is , [C A; T G]^2 leads to the 4x4 matrix (upper right panel), while [C A; T G]^3 generates the 8x8 matrix (lower panel). Please note that the original pattern of 4 nucleotide distribution in 2-D space of the matrices is preserved in the 4x4 and 8x8 matrices [12, 13], just as the double stranded DNA molecules are preserved as a fertilized egg divides to form multi-cellular organs and organisms.

(6) *The Petoukhov coincidence* [14].

Petoukhov generated the 8x8 matrix containing 64 triplets mathematically from the four genetic alphabet (see Figure 3) [12, 13]. A close examination shows that these mathematically generated 64 triplets of nucleotides coincide with the 64 genetic codons as shown in the fourth row of **Figure 1**. The coincidence between the mathematically generated 8x8 matrix of 64 triplets and the evoluionarily generated 8x8 matrix of 64 codons has been referred to as the Petoukhov coincidence in [14] for the convenience of discussions. Since there are 4^3 = 64 possible nucleotide triplets, the number of all possible ways of arranging these triplets into an 8x8 matrix would be 64! which is about 10^89, a gargantuan number, much larger that the age of our universe expressed in seconds, i.e:, 10^17 [12]. Therefore any one of the variants of the codon table such as the one in **Figure 1** carries maximally log_2(10^89) bits of **Shannon information**, or about 300 bits. This may indicate that the protein-coding regions of DNA (which accounts for ~3% of the whole DNA mass in the human genome) can store genetic information up to 300 bits *able to make 300 binary decisions or 300 binary selections.*

(7) In 1993 [15], Trifonov concluded that DNA is a molecular language consisting of *letters*, *words,* and *texts* , carrying *overlapping messages* embodying multiple codes (as in our broadband communication channels of the Internet):

"The classical triplet code is not the only code carried by the sequences. They contain, for example, the gene-splicing code, transcription codes and many other codes. By analyzing a large volume of the nucleotide sequences available, i.e., by performing various computer experiments with the sequences, one can decipher them and extract from them valuable biological information. At the DNA level there are at least two more codes — the DNA shape code and the chromatin code. The overall DNA shape is sequence-dependent and can be described by a set of angles characteristic for various dinucleotide elements — codons of the DNA shape code. The chromatin code provides instructions for histone octamers where along the DNA to form the nucleosomes. This code is expressed as positional periodicity of, primarily, AA and TT dinucleotides. A new RNA code has been described — the translation framing code. The frame seems to be maintained by a synchronizing pattern GCUGCUGCU… hidden in mRNA. Most enigmatic of all is, perhaps, the gene-splicing code. An interesting recent developmentindicates that the gene-splicing pattern in the sequences and the nucleosomal pattern have some common features. This has to do with superposition of the patterns that is characteristic for the sequence language in general which carries simultaneously many codes in one and the same text. This results in an increased complexity of the sequences. Analysis of the protein-coding sequence complexity in eukaryotes and in prokaryotes revealed that the former are simpler. This is interpreted as the result of a spatial separation of the triplet code (carried by exons) and the chromatin code (carried by introns). Perhaps, the necessity of the separation of otherwise conflicting codes is one of the reasons why the intervening sequences had been introduced at all. The nucleotide sequences are writtenin an unbroken manner. One way to detect "words" in such a continuous text is to evaluate the degree of internal correlation by calculating contrast values for the words. This technique allows one to derive vocabularies, which are species- and function-specific. The nucleotide sequences, thus, carry numerous superimposed messages. We do understand only a few of these messages while many more are waiting for their turn to be deciphered."(emphasis added)

Again, for the convenience of discussion, I am suggesting here that we refer to the ability of DNA sequences to encode and transmit multiple *overlapping* messages simultaneously as the "Trifonov mechanism of DNA information storage and transfer (TMDIST)" or more simply the "Trifonov mechanism of DNA language (TMDL)".

(**8**) Independently of Trifonov, I postulated in 1997 that cells use a molecular language that shares 10 out of the 13 design features of the human language [16, 17] as summarized in **Table 2**:

Table 2. The isomorphism between human and cell languages.

	Design features	**Human Language (*Humanese*)**	**Cell Language (*Cellese*)**
1	*Communication channel*	Sound waves Written words	conformational waves concentration waves covalent structures of DNA, etc.
2	*Arbitrariness of signs*	Arbitrary relation between the name and the named, e.g, polysemy	Arbitrary relation between biopolymer linear structures and their 3-dimensional shapes (and hence functions)
3	*Duality* (or *Double articulation*)	1st: Words ---> sentences 2nd: Letters ---> words	1st: 1-D biopolymer --- > 3-D biopolymer 2nd: Monomers ---> 1-D biopolymer
4	*Productivity* (or *Rule-governed creativity*)	Small number of letters, words and grammatical rules can generate an almost infinite number of meaningful sentences	Small number of monomers (20 amino acids, 4 nucleotides) can generate an almost infinite number of functional biopolymers (proteins, DNA, RNA).
5	*Discreteness*	Two word-forms (e.g., bear & pear) are absolutely different	Elements of cell languages (e.g., hormones) are absolutely different because they are molecules with distinct shapes.
6	*Semanticity*	Each word (e.g., salt) has a fixed meaning (i.e., is not sand).	Extracellular messengers trigger specific gene-directed processes.
7	*Displacement*	Talking about things that are remote in space and time	Watson-Crock genes transmitinformation in time and Prigoginian genes transmit information in space.
8	Intercahngeability	Humans can be both senders and receivers of messages	Cells can also be senders and receivers of messages.
9	Complete feedback	Humans can hear and monitor his/her own utterances to check their comprehensibility and correctness of formation as they are produced.	Certain cells in the human body utilize hormones produced by themselves called 'autohormones'. This phenomenon is known as 'autocrine'.
10	Specialization	A signal is said to be 'specialized' if its direct physical effect on the behavior of the receiver are not functionally interrelated.	Hormones or DNA sequences show the phenomenon of specialization because these molecules and their target functions are not materially interrelated.

Of the 10 design features listed in **Table 2**, I have found "Double Articulation" in Row 3 most useful, because it enabled me to distinguish between *letters*, *words*, and *sentences* and identify their molecular counterparts in the cell language (see Row

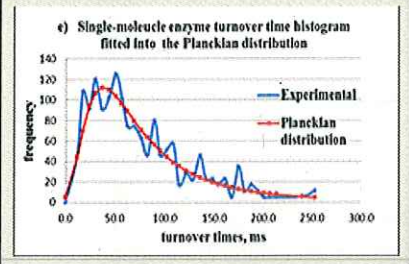
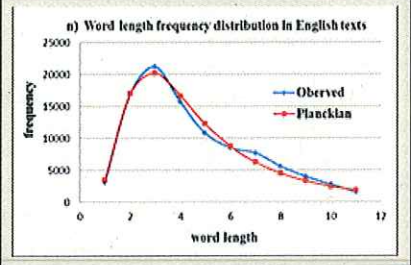
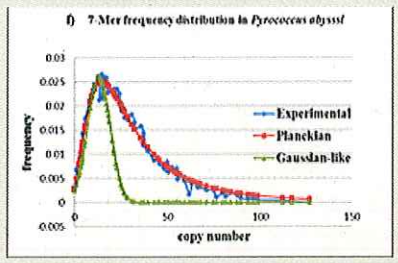
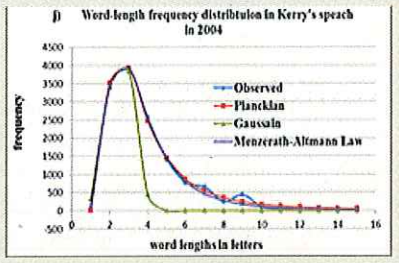
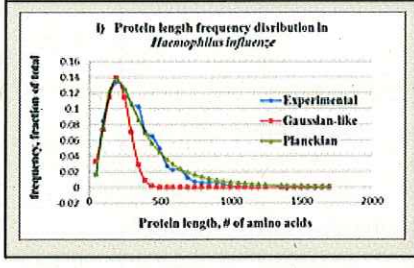
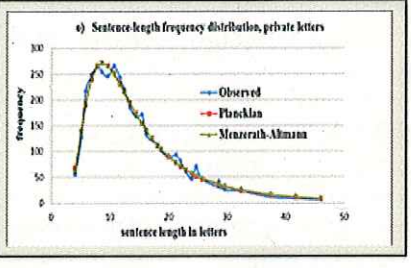
3 in **Table 2**). In 2012 [18], I extended the concept of double articulation to include what I called the "third articulation" which is defined in the last row of the following table (adopted from [18; Section 4.2.2]):

Table 6-3 A *formal* comparison between human and cell languages (Ji 1997a, 1999b).

	Human Language (*Humanese*)	**Cell Language** (*Cellese*)
1. Alphabet (L)	Letters	4 Nucleotides (or 20 Amino acids)
2. Lexicon (W)	Words	Genes (or Polypeptides)
3. Sentences (S)	Strings of words	Sets of genes (or polypeptides) expressed (or synthesized) coordinately in space and time dictated by DNA folds (cell states).
4. Grammar (G)	Rules of sentence formation	The *physical laws* and *biological rules*mapping DNA sequences to folding patterns of DNA (polypeptides) under biological conditions.
5. Phonetics (P)	Physiological structures and processes underlying phonation, audition, and interpretation, etc.	Concentration and mechanical waves responsible for information and energy transfer and transduction driven by *conformons* and *intracellular dissipative structures* (IDSs).
6. Semantics (M)	Meaning of words and sentences	*Codes* mapping molecular signs to gene-directed cell processes
7. First Articulation	Formation of sentences from words	Organization of gene expression events inspace and time through *non-covalent interactions* between DNA and proteins (or Space- and time-dependent non-covalent interactions among proteins, DNA, and RNA molecules). Thus, macromolecular complexes can be viewed as molecular analogs of sentences.
8. Second Articulation	Formation of words from letters	Organization of nucleotides (or amino acids) into genes (or polypeptides) through covalent interactions.
9. Third Articulation	Formation of texts from sentences	Organization of chemical concentration gradients in space and time called *dissipative structures* (Babloyantz 1986, Kondepudi and Prigogine 1998) or *dissipatons* in order to 'reason' and 'compute'.

(**9**) There may be two complementary aspects to the isomorphism between cellese and humanese -- the qualitative and the quantitaive. Table 6-3 lists the qualitative similarity (or qualitative functor, to borrow the category-theoretic idiom) and **Table 3** below lists the quantitative similarities (or quantitative functor). That is, both cellese and humanese obey the same pair mathematical equations -- PDE was derived from physics [4] and the Menzerath-Altmann law was derived from linguistics [19] but they are indistinguishable as far as theri ability to fit long-tailed histograms generated from biology (see the left column in **Table 3**) or linguistics (the right column in **Table 3**). Therefore it seems reasonable to conclude that Tables 6-3 and 3 provide both the qualitative and quantiative evidence to support the validity of the isomorphism between cellese and humanese.

Table 3. The Planckian Distribution Equation (PDE) and the Menzerath-Altmann Law (MAL) connect the *cell* and *human languages*. Adopted from [14].

Cell Language (*Cellese*)	Human Langauge (*Humanese*)
	
	
	

(**10**) Finally, one of the main consequences of combining the recent results of Petoukhov [12, 13] that *genes are coupled via resonances* and my results indicating (i) that *cells use languages isomorphic with the human language* [14, 16, 17] and (ii) that the *wave-particle duality first discovered in physics in the early decades of the 20th century is universal, applying not only to microphysics but also to molecular, cell, and human biology, and beyond* [4] is this:

"It is generally believed that there is only one genetic alphabet consisting of the 4 molecular letters, A, C, G and T. In contrast, the available experimental evidence discussed above indicates
that there are multiple genetic alphabets, each consisting of different number, 4^n, of molecular letters where n is the number of molecular components constituting a molecular letter
and can range from 1 to 5 or more. Thus we may refer to the multiple genetic alphabets postulated here as the 1st-order, the 2nd-order, the 3r-order, . . . , the n^th-order genetic alphabets
(see Table 3, below)"

For the convenience of reference, we may refer to the proposed idea as the "multiple genetic alphabet (MGA) hypothesis". It is hope that the validity of the MGA hypothesis can be tested through the computer-based analysis of currently available genetic data along the lines of research being carried out by Petoukhov [12, 13] and Trifonov [15].

Table 3. Are there more than one genetic alphabets? The structure and function of the cell language inferred on the basis of the postulated isomorphism between human and cell languages [14, 16, 17]. and the role of vibrational resonances in genetic structures [12, 13].

Human Language	**Cell Language**		
	Structure of molecular alphabets		**Function**
Alphabets	1st-order	$4^1 = 4$ singlets (A, C, G, T)	Encoding 1-nucleotide frame shift ?
	2nd-order	$4^2 = 16$ doublets (AC, AG, AT, CA, CG, CT, GA, etc.)	Encoding 2-nucleotide frame shift ? DNA shape code [15], chromatin code [15]

	3rd-order	$4^3 = 64$ triplets (AAA, AAC, AAG, AAT, ACA, etc.)	Encoding amino acids, stop, and start codons
	4th-order	$4^4 = 256$ tetrads [4] (AAAA, AAAC, AAAG, AAAT, etc.)	Translation frame code? [15]?
	5th-order	$4^5 = 1024$ pentads [4] (AAAAA, AAAAC, AAAAG, AAAAT, etc.)	Translation frame code? [15]?
Words	Genes (or the tetra-groups of Petoukhov ?) [13]		Encoding the primary structure of proteins (e.g.,insulin)
Sentences	Gene systems		Encoding protein complexes (e.g., multi-subunit glycolytic enzymes)
Texts	Systems of gene systems		Encoding systems of complexes (e.g., the glycolytic pathway)

Any questions or comments would be welcome as always.

All the best.

Sung

References:

[1] Ramsey, L. and Lee, S. (2016). Our DNA is 99.9% the same as the person sitting next to us — and we're surprisingly similar to a bunch of other living things. https://www.google.com/search?biw=1184&bih=592&tbm=isch&sa=1&ei=LRAOWoKjJIyKmQHq3ZzAAQ&q=Our+DNA+is+99.9%25+the+same+as+the+person+sitting+next+to+us+%E2%80%94+and+we%27re+surprisingly+similar+to+a+bunch+of+other+living+things+Lydia+Ramsey+and+Samantha+Lee+May+6%oq=Our+DNA+is+99.9%25+the+same+as+the+person+sitting+next+to+us+%E2%80%94+and+we%27re+surprisingly+similar+to+a+bunch+of+other+living+things+Lydia+Ramsey+and+Samantha+Lee+May+6%gs_l=psy-ab.3...242.4228.0.8579.4.4.0.0.0.0.0.0..2.0....0...1.1j2.64.psy-ab..2.0.0.0...354.LAXf2f5HScM

[2] Ji, S. (1985). The Bhopalator – A Molecular Model of the Living Cell Based on the Concepts of Conformons and Dissipative Structures. *J. theoret. Biol.* **116**:399-426. PDF at http://www.conformon.net/wp-content/uploads/2012/05/Bhopalator_11.pdf

[3] Cole, S. W. (2014). Human Social Genomics. *PLOS Genetics* **10**(8): e1004601. https://doi.org/10.1371/journal.pgen.1004601.

[4] Ji, S. (2015). Planckian distributions in molecular machines, living cells, and brains: The wave-particle duality in biomedical sciences. In: *Proceedings of the International Conference on Biology and Biomedical Engineering*, Vienna, March 15-17, 2015. Pp. 115-137. http://www.conformon.net/wp-content/uploads/2016/09/PDE_Vienna_2015.pdf

[5] Amazing Resonance Experiment! https://www.youtube.com/watch?v=wvJAgrUBF4w

[6] Patel, K. (2013). Chladni Plates - Sand Vibration Patterns https://physicsbridgecourse.blogspot.com/2013/11/chladni-plates-sand-vibration-patterns.html

[7] Ji, S. (1991). Conformons and Molecular Mchines. In: Molecular Theory of Cell Life and Death, Rutgers University Press, New Brunswick. PDF at http://www.conformon.net/wp-content/uploads/2013/01/thermal_barrier_p31_p51.pdf

[8] Tama, F., Valle, M., Frank, J. and Brooks, C. L. (2003). Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Pro. Nat. Acad. Sci. USA* **100** (16): 9319-9323.

[9] The Nobel Prize in Chemistry 2017. https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/

[10] Herbert, N. (1987). Quantum Reality: Beyond the New Physics, An Excursion into Metaphysics, Anchor Books, Garden City, New York, p. 64.

[11] Ji, S. (1974). Energy and Negentropy in Enzymic Catalysis. *Ann. N. Y. Acad. Sci.* **227**:419-437.

[12] Petoukhov, S. V. (2016). The system-resonance approach in modeling genetic structures. BioSystems 139:1-11. http://www.sciencedirect.com/science/article/pii/S0303264715001732

[13] Petoukihov, S. V. (2017). The rules of long DNA-sequences and tetra-groups of oligonucleotides. https://arxiv.org/ftp/arxiv/papers/1709/1709.04943.pdf

[14] Ji, S. (2017). *The Cell Language Theory; Connecting Mind and Matter.* World Scientific Publications. New Jersey. Chapter 5.

[15] Trifonov, E. N. (1993). http://www.worldscientific.com/doi/abs/10.1142/9789814503655_0009

[16] Ji, S. (1997). Isomorphism between cell and human languages: molecular biological, bioinformatic and linguistic

implications. *BioSystems* **44**: 17-39. PDF at http://www.conformon.net/wp-content/uploads/2012/05/Isomorphism1.pdf

[17] Ji, S. (1999). The Linguistics of DNA: Words, Sentences, Grammar, Phonetics, and Semantics. *Ann. N. Y. Acad. Sci*. **870**:411-417.

[18] Ji, S. (2012). *Molecular Theory of the Living Cell: Concepts, Molecular Mechanisms, and Biomedical Applications*. Springer, New York. P. 166.

[19] Menzerath' law. https://en.wikipedia.org/wiki/Menzerath%27s_law

Menzerath's law - Wikipedia

en.wikipedia.org

Menzerath's law, or Menzerath–Altmann law (named after Paul Menzerath and Gabriel Altmann), is a linguistic law according to which the increase of a linguistic ...